制約付き最適化問題に対する主双対空間での正確な内点メリット関数(科学技術における数値計算の理論と応用)

山下, 浩

数理解析研究所講究録
944 巻 1996 年 59-67

制約付き最適化問題に対する主双対空間での正確な内点メリット関数

山下浩 (Hiroshi Yamashita)
(株) 数理システム

Abstract

本論文では内点法による非線形最適化問題の解法を扱う．本方法は修正 KKT 条件に対する主双対空間上での Newton 法を基礎としているが，バリヤパラメータが主双対変数に依存している点で通常の方法と異なっている．このバリヤパラメータ関数を利用して主双対空間での正確なメリット関数を定義することが出来る．そして，適当な条件の下で Newton 法の反復ベクトルがこのメリット関数に対する降下方向になっていることが示される．主双対空間で Armijo タイプの直線探索を実行することによって大域的収束性をもつアルゴリズムをつくることが可能であることを示す．

1 はじめに

本論文で扱う問題は

$$\begin{array}{lll} \text{最小化} & f(x), & x \in \mathbf{R}^n, \\ \text{条　件} & g(x)=0, x \geq 0 & \end{array} \tag{1}$$

である．ここで $f:\mathbf{R}^n \to \mathbf{R}$ と $g:\mathbf{R}^n \to \mathbf{R}^m$ は 2 回連続微分可能であるとする．線形計画問題に対する内点法は理論的にも実際的にも非常に有効であることが知られている．([2],[4],[6]) しかし，非線形計画問題に対する同様な試みは現在のところそれほど多くはない．([1],[3],[5],[7],[9],[10],[8]) 本論文では [9] で提案された方法を拡張して，主双対空間上で大域的収束性をもった内点法を考える．

上記の問題に対するラグランジュ関数を

$$L(w)=f(x)-y^tg(x)-z^tx$$

と定義する．ここで $y \in \mathbf{R}^m$ と $z \in \mathbf{R}^n$ はそれぞれ等式制約条件と非負条件に対する双対変数である．また，$w=(x,y,z)^t$ である．上記の問題の Karush-Kuhn-Tucker 条件は

$$r(w) \equiv \begin{pmatrix} r_L(w) \\ r_E(w) \\ r_C(w) \end{pmatrix} \equiv \begin{pmatrix} \nabla L_x(w) \\ g(x) \\ Xz \end{pmatrix} = 0, x \geq 0, z \geq 0$$

と表わされる．ここで $r(w)$ は等式条件に対する残差ベクトルで，Xは行列 $\mathrm{diag}(x_1, \cdots, x_n)$ を表わす．

μを正の定数，e をすべて 1 からなる n 次元ベクトルとして，[9] では修正 KKT 条件

$$\tilde{r}(w,\mu) \equiv r(w) - \mu\hat{e} = 0, \hat{e} = \begin{pmatrix} 0 \\ 0 \\ e \end{pmatrix} \tag{2}$$

を Newton 法：

$$J(w)\Delta w \equiv \begin{pmatrix} \nabla_x^2 L(w) & A(x)^t & -I \\ A(x) & 0 & 0 \\ Z & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = -r(w) + \mu\hat{e}$$

によって解くアルゴリズムが提案された．ここで，$J(w) \in \mathbf{R}^{(2n+m)\times(2n+m)}$は関数 $r(w)$ のヤコビ行列で，$A(x) = \nabla g(x)^t$ ，$Z = \mathrm{diag}(z_1, \cdots z_n)$ である．また，$\Delta w = (\Delta x, \Delta y, \Delta z)^t$ は反復ベクトルである．アルゴリズムの大域的収束のためにバリヤ-ペナルティ関数

$$F(x) = f(x) - \mu \sum_i \log(x_i) + \rho \sum_i |g_i(x)| \tag{3}$$

を導入して，適当な直線探索を採用することによって，$\mu > 0$ に対して (2) の解として定義されるセンタへの大域的収束が証明された．そして，μの減少列に対して近似的最小化が実施されて問題 (1) の近似的 KKT 点が求まる．[9] では，変数 x のみに対するメリット関数 (3) を利用するために主変数 x と双対変数 yと zは異なる扱いを受けた．本論文では変数 wに対するメリット関数を使用する内点法を提案する．この意味で主変数と双対変数は同等に扱われることになる．本方法は内点法であるので，生成される点列は

$$S_0 = \{w \in \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^n \,|x > 0, z > 0\}$$

で定義される集合 S_0に存在する．以下では，固定したバリヤパラメータ$\mu > 0$ を使用した残差ベクトル (2) のかわりに，バリヤパラメータ関数$\mu : \mathbf{S}_0 \to \mathbf{R}^1$を使用した以下の修正 KKT 条件を扱う：

$$\tilde{r}(w, \mu(w)) \equiv r(w) - \mu(w)\hat{e} = 0 \tag{4}$$

この条件を修正 Newton 法：

$$J(w)\Delta w = -r(w) + \mu(w)\hat{e} \tag{5}$$

によって解く方法が本論文のアルゴリズムの基礎となる．

2 主双対バリヤペナルティ関数

まず，S_0で定義される主双対バリヤペナルティ関数を

$$F(w) = f(x) - \mu(w) \sum_i \log\left(\frac{x_i}{\overline{x}_i}\right) + \rho \sum_i |g_i(x)| \tag{6}$$

とする．ここで，$\bar{x}_i\,(i=1,\cdots,n)$ は$-\log(x_i/\bar{x}_i)$ が常に正となるような十分大きい正数とする．

ベクトル $s=(s_x,s_y,s_z)$ を wの空間でのステップとする．このとき，メリット関数 $F(w+s)$ の 1 次近似 $F_l(w,s)$ を

$$\begin{aligned} F_l(w,s) &= F(w)+\nabla f(x)^t s_x-\mu(w)e^tX^{-1}s_x+\rho\sum_i\left|g_i(x)+\nabla g_i(x)^t s_x\right|-\rho\sum_i|g_i(x)| \\ &\quad -\Delta\mu(w,s)\sum_i\log\left(\frac{x_i}{\bar{x}_i}\right) \end{aligned}$$

によって定義する．ここで，$\Delta\mu(w,s)$ は$\mu(w+s)-\mu(w)$ に対する 1 次近似であり，以下で定義される．そして，$F(w)$ の変化分の 1 次近似を

$$\Delta F_l(w,s)=F_l(w,s)-F(w)$$

によって定義する．修正 Newton 方程式 (5) の解をΔwとして $s=\Delta w$とおくと

$$\begin{aligned} \Delta F_l(w,\Delta w) &= \nabla f(x)^t\Delta x-\mu(w)e^tX^{-1}\Delta x+\rho\sum_i\left|g_i(x)+\nabla g_i(x)^t\Delta x\right| \\ &\quad -\rho\sum_i|g_i(x)|-\Delta\mu(w,\Delta w)\sum_i\log\left(\frac{x_i}{\bar{x}_i}\right) \\ &\le -\Delta x^t(\nabla_x^2L(w)+X^{-1}Z)\Delta x-(\rho-\|y+\Delta y\|_1)\sum_i|g_i(x)| \\ &\quad -\Delta\mu(w,\Delta w)\sum_i\log\left(\frac{x_i}{\bar{x}_i}\right) \end{aligned} \tag{7}$$

が成立する．この不等式の導出は [9] と同様になされる．

本論文の主要な提案となる S_0上のバリヤペナルティ関数を次のように定義する．

$$\mu(w)=\frac{\kappa\|r(w)\|_1^{p+q}}{\left(\prod_{i=1}^n x_iz_i\right)^{p/n}} \tag{8}$$

ここで$\kappa\in(0,1)$，$p>0$，$q>0$ は定数である[1]．$\Delta\mu(w,\Delta w)$ を計算するためには $\|r(w)\|_1$ の変化の 1 次近似（それを$\Delta\|r(w)\|_1$と表わす）を $r(w)$ 自身の変化の 1 次近似

$$\Delta r(w,\Delta w)\equiv J(w)\Delta w=-r(w)+\mu(w)\hat{e}$$

から計算しなくてはならない．この量は

$$\begin{aligned} \Delta\|r(w)\|_1 &\equiv \|r(w)+\Delta r(w,\Delta w)\|_1-\|r(w)\|_1 \\ &= \|\mu(w)\hat{e}\|_1-\|r(w)\|_1 \\ &= -\|r(w)\|_1+n\mu(w) \end{aligned}$$

[1](8) においては l_1ノルムが採用されている．しかし，l_2あるいは l_∞ノルムも同様に使用できる．

によって計算される．これらの量を利用してΔwに対する$\mu(w)$の変化の１次近似を

$$\Delta\mu(w,\Delta w) \equiv \frac{\kappa(p+q)\left\|r(w)\right\|_1^{p+q-1}\Delta\left\|r(w)\right\|_1}{\left(\prod_{i=1}^{n} x_i z_i\right)^{p/n}} - \frac{\kappa p\left\|r(w)\right\|_1^{p+q}\Delta\left(\prod x_i z_i\right)}{n\left(\prod_{i=1}^{n} x_i z_i\right)^{p/n+1}} \tag{9}$$

と表わすことができる．

以下の補助定理が本論文の中心的結果である．

Lemma 1 $w \in S_0$でΔwは *(5)* をみたすとする．このとき

$$\Delta\mu(w,\Delta w) \le -q(1-\kappa n^{p+1}(1+q/p)^p\left\|r(w)\right\|_1^{q-1})\mu(w) \tag{10}$$

が成立する．

証明 （9）から

$$\begin{aligned}
\Delta\mu(w,\Delta w) &= \frac{\kappa(p+q)\left\|r(w)\right\|_1^{p+q-1}\left\{-\left\|r(w)\right\|_1+n\mu(w)\right\}}{\left(\prod_{i=1}^{n} x_i z_i\right)^{p/n}} - \frac{\kappa p\left\|r(w)\right\|_1^{p+q}\sum_{i=1}^{n}\frac{(\mu(w)-x_i z_i)}{x_i z_i}}{n\left(\prod x_i z_i\right)^{p/n}} \\
&= \frac{\kappa\left\|r(w)\right\|_1^{p+q}}{\left(\prod_{i=1}^{n} x_i z_i\right)^{p/n}}\left\{-(p+q)+\frac{n(p+q)\mu(w)}{\left\|r(w)\right\|_1}-\sum_{i=1}^{n}\left(\frac{p\mu(w)}{n x_i z_i}-\frac{p}{n}\right)\right\} \\
&= \frac{\kappa\left\|r(w)\right\|_1^{p+q}}{\left(\prod_{i=1}^{n} x_i z_i\right)^{p/n}}\left\{-q+\mu(w)\left(\frac{(p+q)n}{\left\|r(w)\right\|_1}-\sum_{i=1}^{n}\frac{p}{n x_i z_i}\right)\right\} \\
&= -q\mu(w)+\frac{\kappa\left\|r(w)\right\|_1^{p+q}\mu(w)}{\left(\prod_{i=1}^{n} x_i z_i\right)^{p/n}}\left(\frac{(p+q)n}{\left\|r(w)\right\|_1}-\sum_{i=1}^{n}\frac{p}{n x_i z_i}\right) \qquad (11)
\end{aligned}$$

となる．

もし

$$\frac{(p+q)n}{\left\|r(w)\right\|_1}-\sum_{i=1}^{n}\frac{p}{n x_i z_i} \le 0$$

ならば（11）から

$$\Delta\mu(w,\Delta w) \le -q\mu(w) \tag{12}$$

となる．

そこで，

$$\frac{(p+q)n}{\left\|r(w)\right\|_1}-\sum_{i=1}^{n}\frac{p}{n x_i z_i} > 0 \tag{13}$$

が成立している場合を考える．以下では，良く知られた関係

$$\frac{x^t z}{n} \geq \left(\prod_{i=1}^{n} x_i z_i\right)^{1/n}$$

と

$$\sum_{i=1}^{n} \frac{1}{n x_i z_i} \geq \frac{1}{\left(\prod_{i=1}^{n} x_i z_i\right)^{1/n}} \tag{14}$$

を使用する．まず

$$\begin{aligned}
\frac{(p+q)n}{\|r(w)\|_1} - \sum_{i=1}^{n} \frac{p}{n x_i z_i} &\leq \frac{(p+q)n}{\|r(w)\|_1} - \frac{p}{\left(\prod_{i=1}^{n} x_i z_i\right)^{1/n}} \\
&\leq \frac{qn}{\|r(w)\|_1} + \frac{pn}{x^t z} - \frac{p}{\left(\prod_{i=1}^{n} x_i z_i\right)^{1/n}} \\
&\leq \frac{qn}{\|r(w)\|_1}
\end{aligned}$$

であるから (11) より

$$\Delta\mu(w, \Delta w) \leq -q\mu(w) + \frac{\kappa q n \mu(w) \|r(w)\|_1^{p+q-1}}{\left(\prod_{i=1}^{n} x_i z_i\right)^{p/n}} \tag{15}$$

となる．そして, (13) と (14) から

$$\frac{(p+q)n}{\|r(w)\|_1} > \sum_{i=1}^{n} \frac{p}{n x_i z_i} \geq \frac{p}{\left(\prod_{i=1}^{n} x_i z_i\right)^{1/n}},$$

となり

$$\frac{\|r(w)\|_1}{\left(\prod_{i=1}^{n} x_i z_i\right)^{1/n}} < \frac{(p+q)n}{p} \tag{16}$$

を得る. (15) と (16) から

$$\begin{aligned}
\Delta\mu(w, \Delta w) &\leq -q\mu(w) + \kappa q n^{p+1} (1 + q/p)^p \|r(w)\|_1^{q-1} \mu(w) \\
&= -q(1 - \kappa n^{p+1} (1 + q/p)^p \|r(w)\|_1^{q-1}) \mu(w)
\end{aligned}$$

となり補助定理が証明された． □

この補助定理から，もし定数$\lambda \in (0,1)$が存在して

$$\kappa n^{p+1} (1 + q/p)^p \|r(w)\|_1^{q-1} \leq \lambda < 1 \tag{17}$$

ならば

$$\Delta\mu(w) \le -q(1-\lambda)\mu(w) < 0$$

となる．もし $q=1$ ならば条件 (17) は

$$\kappa n^{p+1}(1+1/p)^p < 1$$

となり定数のみを含む．そして補助定理 1の不等式 (10) は

$$\Delta\mu(w) \le -(1-\kappa n^{p+1}(1+1/p)^p)\mu(w) < 0$$

となる．最も単純な p と qの選びかたは

$$p=q=1$$

である．この場合，上記の関係は

$$2\kappa n^2 < 1$$

と

$$\Delta\mu(w) \le -(1-2\kappa n^2)\mu(w) < 0$$

となる．

3 アルゴリズムとその収束性

前節では，適当な条件がみたされれば (7) と補助定理 1 から (5) によって計算される方向Δwはメリット関数 $F(w)$ の降下方向になることが示された．したがって，Δw方向に直線探索を行うことによって関数 $F(w)$ の値を減少させることができる．直線探索におけるステップ巾の決定のために Armijo のルールを使用する．まず，許容領域の境界までの最大のステップとして

$$\alpha_{\max} = \min\left\{\min_i\left\{\frac{-x_i}{\Delta x_i}\middle|\Delta x_i < 0\right\}, \min_i\left\{\frac{-z_i}{\Delta z_i}\middle|\Delta z_i < 0\right\}\right\}$$

を計算する．すなわち$\alpha \in [0, \alpha_{\max})$ は S_0における内点を与える．次の点へのステップは

$$\alpha = \bar{\alpha}\beta^l, \quad \bar{\alpha} = \min\{\gamma\alpha_{\max}, 1\}$$

によって与えられる．ここで，$\gamma \in (0,1)$ と$\beta \in (0,1)$ は定数で，lは

$$F(w+\bar{\alpha}\beta^l\Delta w) - F(w) \le \varepsilon_0\bar{\alpha}\beta^l\Delta F_l(w,\Delta w)$$

をみたす最小の正整数である．ここで$\varepsilon_0 \in (0,1)$ である．

以下に記述されるアルゴリズムの収束を証明するために，関数 $F(w)$ の方向 s に沿った方向微分 $F'(w,s)$ が必要となる．

$$F'(w,s) = \lim_{\alpha\downarrow 0}\frac{F(w+\alpha s) - F(w)}{\alpha}$$

以下の二つの補助定理は [9] と同様に証明される．

Lemma 2 $w \in S_0$, $w+s \in S_0$とする．このとき

$$F(w) + F'(w,s) \le F_l(w,s)$$

が成立し，$\theta \in (0,1)$ が存在して

$$F(w+s) \le F(w) + F'(w+\theta s, s)$$

となる． □

Lemma 3 $w \in S_0$, $w+s \in S_0$と$\varepsilon_0 \in (0,1)$ が与えられたとき，$F_l(w,s) < 0$ ならば十分小さな$\alpha > 0$ に対して

$$F(w+\alpha s) - F(w) \le \varepsilon_0 \alpha F_l(w,s)$$

が成立する． □

本論文のアルゴリズムを以下に示す．

Algorithm

(初期化) $w_0 \in S_0$, $\rho > 0$, $\gamma \in (0,1)$, $\beta \in (0,1)$, $\varepsilon_0 \in (0,1)$ とする．$\varepsilon \ge 0$ を収束判定パラメータとする．$k=0$ とおく．

(Iteration) while $(\|r(w_k)\| > \varepsilon)$ repeat {

(5) によってベクトルΔw_kを計算する;

ステップ巾の初期試行値$\bar{\alpha}_k$を

$$\begin{aligned} \alpha_{k\max} &= \min\left\{\min_i\left\{\frac{-(x_k)_i}{(\Delta x_k)_i}\middle|(\Delta x_k)_i < 0\right\}, \min_i\left\{\frac{-(z_k)_i}{(\Delta z_k)_i}\middle|(\Delta z_k)_i < 0\right\}\right\} \\ \bar{\alpha}_k &= \min\{\gamma\alpha_{k\max}, 1\} \end{aligned}$$

によって計算する;

$$F(w_k + \bar{\alpha}_k\beta^{l_k}\Delta w_k) - F(w_k) \le \varepsilon_0\bar{\alpha}_k\beta^{l_k}\Delta F_l(w_k, \Delta w_k);$$

をみたす最小の正整数 l_kを求める;

$$\alpha_k = \bar{\alpha}_k\beta^{l_k};$$

$$w_{k+1} = w_k + \alpha_k\Delta w_k;$$

とおく；

} □

以下の定理は上記のアルゴリズムの大域的収束性を証明する．

Theorem 1 点 w_0における関数 $F(w)$ の準位集合$\Lambda_F(w_0)$ が有界で，定数 $0 < \lambda < 1$ と $0 < \delta < 1$ が存在して任意の $w \in \Lambda_F(w_0)$ において *(17)* と $x_i < \delta\bar{x}_i, i = 1, \cdots n$ が成立しているとする．さらに準位集合上でΔwは一様に有界，それぞれの kに対して $\nabla^2 L(w_k)$ は非負定値で，$\rho \ge \|y_k + \Delta y_k\|_1$が成立しているとする．このとき，点列 $\{w_k\}$ の任意の集積点は KKT 点である．

証明. 仮定と (7), (17) より

$$\begin{aligned} F(w_{k+1}) - F(w_k) &\le \varepsilon_0 \alpha_k \Delta F_l(w_k, \Delta w_k) \\ &\le \varepsilon_0 \alpha_k \left\{ -\Delta x^t (\nabla_x^2 L(w) + X^{-1} Z)\Delta x - (\rho - \|y + \Delta y\|_1) \sum_i |g_i(x)| \right. \\ &\qquad \left. -\Delta \mu(w, \Delta w) \sum_i \log \left(\frac{x_i}{\overline{x}_i} \right) \right\} \\ &\le -\varepsilon_0 \alpha_k \Delta \mu(w, \Delta w) n \log \delta \\ &\le -\varepsilon_0 \alpha_k q n (1 - \lambda) \mu(w) \log \delta < 0 \end{aligned}$$

を得る．列 $\{F(w_k)\}$ は狭義単調減少で下に有界だから

$$\lim_{k \to \infty} \alpha_k \mu(w_k) = 0$$

となる．もし $\liminf_{k \to \infty} \alpha_k > 0$ ならば$\mu(w_k) \to 0$ となり定理は証明される．そこで，部分列 $K \subset \{0, 1, 2, \cdots\}$ が存在して $l_k \to \infty, k \in K$となる場合を考える．このとき一般性を失うことなく十分大きな $k \in K$に対して $l_k > 0$ であると仮定できる．$l_k > 0$ ならば，点 $w_k + \alpha_k \Delta w_k / \beta$は不等式

$$F(w_k + \alpha_k \Delta w_k / \beta) - F(w_k) > \varepsilon_0 \alpha_k \Delta F_l(w_k, \Delta w_k) / \beta \tag{18}$$

をみたす．したがって，補助定理 2より$\theta_k \in (0, 1)$ が存在して

$$\begin{aligned} F(w_k + \alpha_k \Delta w_k / \beta) - F(w_k) &\le \alpha_k F'(w_k + \theta_k \alpha_k \Delta w_k / \beta, \Delta w_k) / \beta \\ &\le \alpha_k \Delta F_l(w_k + \theta_k \alpha_k \Delta w_k / \beta, \Delta w_k) / \beta \end{aligned} \tag{19}$$

となる．このとき，(18) と (19) から

$$\varepsilon_0 \Delta F_l(w_k, \Delta w_k) < \Delta F_l(w_k + \theta_k \alpha_k \Delta w_k / \beta, \Delta w_k)$$

を得る．この不等式から

$$F_l(w_k + \theta_k \alpha_k \Delta w_k / \beta, \Delta w_k) - F_l(w_k, \Delta w_k) > (\varepsilon_0 - 1) \Delta F_l(w_k, \Delta w_k) > 0$$

が導かれる．$l_k \to \infty, k \in K$ならば上の不等式の左辺は 0 に収束するから$\Delta F_l(w_k, \Delta w_k) \to 0, k \in K$となる．このことと補助定理 1から$\mu(w_k) \to 0, k \in K$である．したがって，定理は証明された． □

参考文献

[1] A.S.El-Bakry, R.A.Tapia, T.Tsuchiya and Y.Zhang, *On the formulation and theory of the primal-dual Newton interior-point method for nonlinear programming*, TR92-40, Dept. of Computational and Applied Mathematics, Rice University, Texas, USA, December 1992 (revised October 1993). R.Fontecilla, Inexact secant methods for nonlinear constrained optimization,

[2] M.Kojima, S.Mizuno and A.Yoshise, A primal-dual interior-point algorithm for linear programming, in *Progress in Mathematical Programming, Interior-Point and Related Methods*, N.Megiddo ed., Springer-Verlag, New York, 1989, pp.29-47.

[3] L.S.Lasdon, J.Plummer and G.Yu, *Primal-dual and primal interior point algorithms for general nonlinear programs*, Manuscript, Texas, USA, February 1993.

[4] I.J.Lustig, R.E.Marsten and D.F.Shanno, Computational experience with a primal-dual interior point method for linear programming, *Linear Algebra and its Applications,* 152 (1991), pp.191-222.

[5] G.P.McCormick, *Resolving the shell dual with a nonlinear primal-dual algorithm*, Report GWU/OR/Serial T-545/91, Dept. of Operations Research, The George Washington University, Washington DC, USA, February 1991.

[6] K.A.McShane, C.L.Monma and D.F.Shanno, An implementation of a primal-dual interior point method for linear programming, *ORSA Journal on Computing*, 1 (1989) pp.70-83.

[7] J.-P.Vial, *Computational experience with a primal-dual interior-point method for smooth convex programming*, Dept. of Commercial and Industrial Economics, University of Geneva, Geneva, Switzerland, August 1992.

[8] H.Yabe and H.Yamashita, On the rate of convergence of primal-dual interior point methods based on quasi-Newton methods, in *The State of the Art of Scientific Computing and its Prospect*, Research Report (RIMS Kokyuroku) No.880, Research Institute for Mathematical Sciences, Kyoto University, Kyoto, Japan, in Japanese, July (1994), pp.211-219.

[9] H.Yamashita, *A globally convergent primal-dual interior point method for constrained optimization,* Technical Report, Mathematical Systems Institute Inc., Tokyo, Japan, April 1992 (revised March 1994).

[10] H.Yamashita and H.Yabe, *Superlinear and quadratic convergence of some primal-dual interior point methods for constrained optimization*, Technical Report, Mathematical Systems Institute Inc., Tokyo, Japan, June 1993 (revised November 1994).